# Characterization of the highly cited articles published by a genetics research department: an exploratory study

**Mirian Ramirez**, *MLIS* – Research Metrics Librarian, Ruth Lilly Medical Library

**Elizabeth C. Whipple**, *MLS, AHIP,* Assistant Director of Research & Translational Sciences, Ruth Lilly Medical Library

## BACKGROUND AND OBJECTIVE

To identify and assess highly cited papers included in the *Web of Science Essential Science Indicators* (ESI), this study looked at authors affiliated with a genetics department at a School of Medicine from 2010-2019. For this study, we examined authorship characteristics, including female authorship trends, subject trends, and intramural and extramural co-authorship networks. This approach aims to highlight research impact trends to inform the department's leaders in decision-making for future publication and research strategy development directions.

### MOLECULAR BIOLOGY & GENETICS RESEARCH FIELD

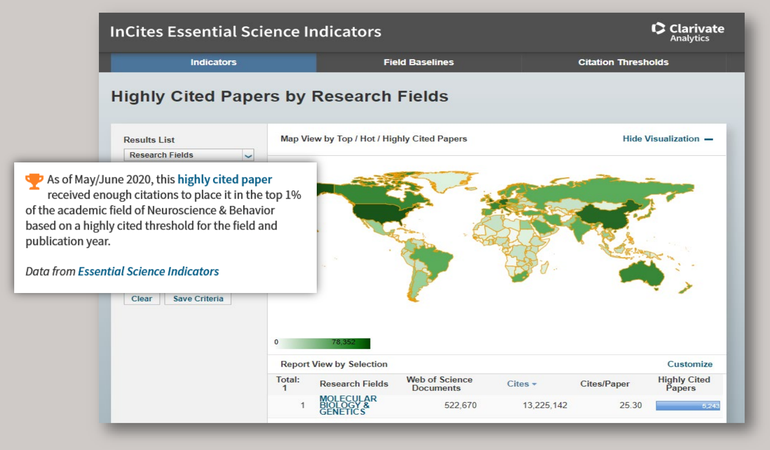
- We selected this research field as a pilot; this ESI research field ranked No. 1 (out of 22 fields) with the highest average citations per article.

| | Research Fields | Web of Science Documents | Cites | Cites/Paper | Highly Cited Papers |
|---|---|---|---|---|---|
| 1 | MOLECULAR BIOLOGY & GENETICS | 531,700 | 13,731,662 | 25.83 | 5,321 |
| 2 | IMMUNOLOGY | 290,965 | 5,960,905 | 20.49 | 2,905 |
| 3 | SPACE SCIENCE | 163,323 | 3,228,064 | 19.76 | 1,644 |
| 4 | NEUROSCIENCE & BEHAVIOR | 573,970 | 11,312,626 | 19.71 | 5,748 |
| 5 | MULTIDISCIPLINARY | 25,033 | 487,884 | 19.49 | 250 |
| 6 | BIOLOGY & BIOCHEMISTRY | 825,186 | 15,474,261 | 18.75 | 8,317 |
| 7 | MICROBIOLOGY | 237,695 | 4,117,715 | 17.32 | 2,395 |
| 8 | MATERIALS SCIENCE | 1,064,555 | 18,320,501 | 17.21 | 10,635 |
| 9 | CHEMISTRY | 1,938,657 | 32,789,220 | 16.91 | 19,299 |
| 10 | GEOSCIENCES | 540,008 | 7,871,253 | 14.58 | 5,469 |

*Source: Essential Science Indicators database*

### ESSENTIAL SCIENCE INDICATORS (ESI):

- Tool integrated to the *Web of Knowledge* platform (by Clarivate analytics).
- ESI analyzes **articles** and **reviews** from the *Web of Science* Core collection journals to determine top-performing articles, authors, institutions, countries/regions, and journals across 22 broad disciplines.
- Tracks ”highly cited” articles that are the top one percent in each of the 22 ESI subject areas per year. They are based on the **most recent 10 years of publications**.




### ESI “HIGHLY CITED” PAPERS

- papers that perform in the top 1% based on the number of **citations** received when compared to other papers published in the same field in the same year.
- Methodology to select Highly Cited Papers – those published and cited during 2010-2019 and which then ranked in the top 1% by citations for their ESI field and year.

> *“Highly Cited Papers are papers that have received enough citations to place them in the top 1% when compared to all other papers published in the same year in the same field, i.e. 2008 Physics papers are only compared to other 2008 Physics papers to determine whether they have been cited enough to rank in the top 1%.”*[1]

## MATERIALS AND METHODS

**Data Source and Search Strategy**

To collect the data we performed a three-phase approach. The searches were conducted in June 2020.

- Phase 1 | Scopus® | find documents affiliated with the genetics department
- Phase 2 | Clarivate Web of Science™ | queried the WoS for all the citations resulting in the *Scopus* search and limited to "Highly Cited” papers. JIF collected from the *Journal Citation Reports*.
- Phase 3 | PubMed | compiled MeSH terms and funding of highly cited articles identified

**Materials**

We used the *OpenRefine* tool to perform cleanup, *Excel* to work with datasets of bibliometric data, and *VOSviewer* software to construct the network visualizations.

## RESULTS AND CONCLUSIONS

For the 2010-2019 time frame, a total of 1,077 articles (original and review articles) were published by this department, with 37 documents categorized as Highly Cited Papers. Identified documents gathered a total of 13,239 citations in total (average of 357.81 citations per article). The majority of articles were case-control studies, systematic reviews (SRs), and/or meta-analyses (70%). The average number of authors for these highly cited articles was 144, with 468 as the largest number and the smallest 2 authors. The highly cited articles were published in 17 high-impact journals, ranked in Q1 and Q2, from diverse WoS subject categories, including Neuroscience, Oncology, and Genetics research fields. Indicative of that, the highly cited papers have a strong relationship between the impact factors. Based on an analysis of MeSH terms, the main subject areas focus on "*Genome-wide Association Study*", "*Genetic predisposition to disease*," and *Polymorphism, single nucleotide*" (most used MeSH descriptors).

### MAIN INFORMATION OF THE COLLECTION

| | | |
|---|---|---|
| **37** Highly cited articles | **3.4%** **A**rticles “highly cited” | **357.81** Avg citations per article |
| **17** Journals | **144** Avg authors per article | **2,632** Unique authors |

**Figure 1. Distribution and evolution of articles from 2010-2019**. A total of 37 (of 1077) articles were identified as highly cited articles, with the majority being original articles (34 of 37), followed by review articles (3 of 37)

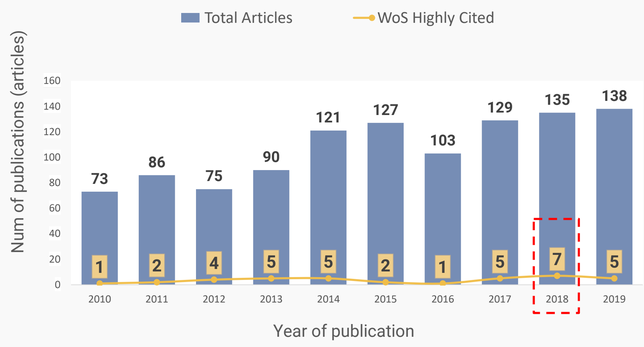
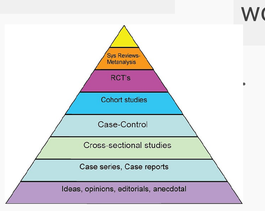



**Figure 2. Study design/EBM levels of evidence.** The majority of articles were case-control studies and SR and/or meta-analysis (26 of 37, 70%).

| | Total articles | Total citations | Average citations |
|---|---|---|---|
| Systematic Reviews & Meta-Analysis | 12 | 5628 | 469 |
| Randomized Control Trials (RCTs) | 1 | 89 | 89 |
| Cohort Studies | 4 | 1326 | 331.5 |
| Case-Control Studies | 14 | 4890 | 349.29 |
| Case Report or Case Series, Narrative Reviews | 6 | 1306 | 217.67 |




*Evidence Pyramid*

**Figure 3. Authorship –Women authors distribution**. Out of 61 IUSM authors, 19 are women authors. Out of 27 authors affiliated with the genetics department, 9 are women authors.

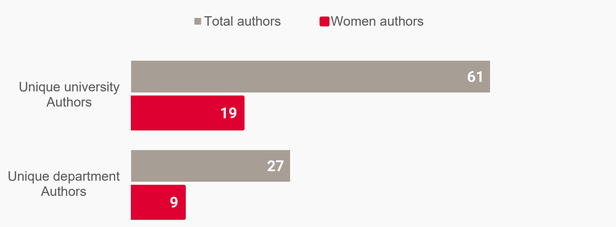



**Figure 4. Top 5 journals publishing highly cited articles (2010-2019).** Highly cited articles were published in 17 high-impact journals ranked in Q1 and Q2. Most of them were published in *Nature Genetics*.

| Source Title | No. of articles | Quartile In Category | Journal Impact Factor | Citation count | Average citation per article |
|---|---|---|---|---|---|
| NATURE GENETICS | 11 | Q1 | 27.605 | 7420 | 674.55 |
| MOLECULAR PSYCHIATRY | 4 | Q1 | 12.384 | 539 | 134.75 |
| NATURE | 3 | Q1 | 42.779 | 1559 | 519.67 |
| NATURE NEUROSCIENCE | 3 | Q1 | 20.071 | 567 | 189 |
| ALZHEIMERS & DEMENTIA | 2 | Q1 | 17.127 | 107 | 53.50 |

**Figure 5.Top 10 funding agencies of highly cited articles.** 713 funding agencies identified (Source: *Web of Science* Results Analysis).

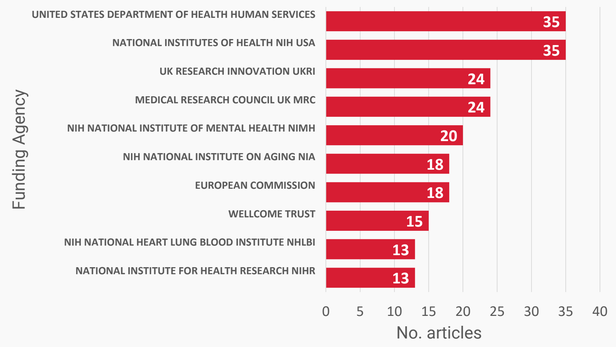



**Figure 6. Intramural collaboration network** of highly cited articles (2010-2019).

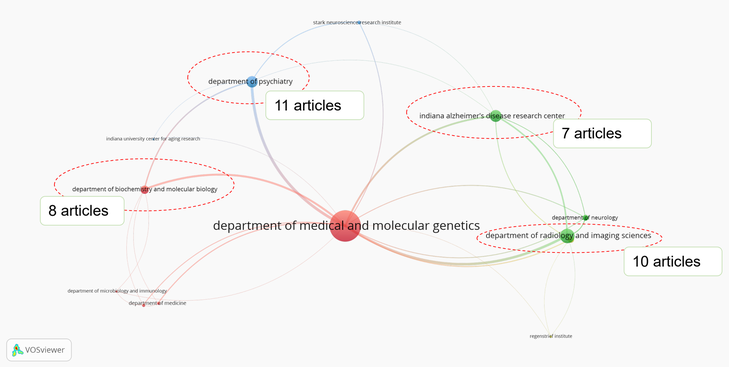
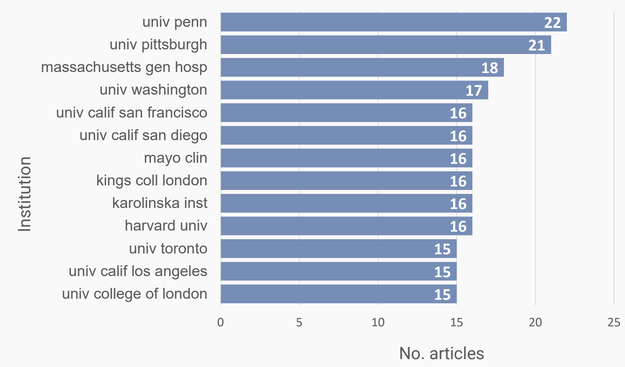



**Figure 7. Collaboration with external institutions.** Top institutions that co-authors at least 20 publications from the IUSM M&MG papers. (source of the analysis: *VOSviewer*)




## REFERENCES

1. Faeth K. LibGuides: Essential Science Indicators: Learn the Basics [Internet]. Essential Science Indicators: Learn the Basics. 2021 [cited 2021 April 15]. Available from: https://clarivate.libguides.com/c.php?g=593878&p=4107958
2. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer’s disease. Nat Genet. 2013 Dec;45(12):1452–8.